

B6

[19] 中华人民共和国国家知识产权局

## [12] 发明专利申请公开说明书

[21] 申请号 99124253.X

[51] Int. Cl<sup>7</sup>

C07K 14/00  
C07K 14/435 C07K 16/00  
C07K 16/18 C07H 21/00  
C12N 15/11 C12N 15/11  
C12N 15/63 C12N 15/64  
C12P 21/00 C01N 33/68  
C12Q 1/68 A61K 38/16  
[11] 公开号 CN 1299825A

[43] 公开日 2001 年 6 月 20 日

[22] 申请日 1999.12.16 [21] 申请号 99124253.X  
[71] 申请人 上海生元基因开发有限公司  
地址 200001 上海市北京东路 668 号 6 楼  
[72] 发明人 毛裕民 谢毅

[74] 专利代理机构 上海市华诚律师事务所  
代理人 徐申民

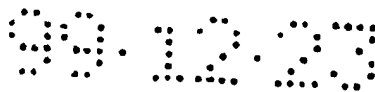
权利要求书 2 页 说明书 18 页 附图页数 1 页

[54] 发明名称 新的人锌指转录因子及其编码序列

[57] 摘要

本发明公开了一种新的人锌指转录因子 (Novel Human Zinc Finger Transcription Factor, 简称“BioZTF45”), 编码此多肽的多核苷酸和经重组技术产生这种多肽的方法。本发明还公开了将此多肽和多核苷酸用于治疗神经系统疾病、血液性恶性疾病、发育紊乱症、肿瘤、内分泌系统疾病、免疫系统疾病等疾病的方法。本发明还公开了抗此多肽的拮抗剂及其治疗作用。本发明还公开了基于鉴别此核酸序列中的突变和此多肽表达水平改变的诊断测定方法。本发明还公开了编码这种新的 BioZTF45 的多核苷酸的用途。

ISSN 1008-4274



## 权 利 要 求 书

---

1、一种分离的人 BioZFTF45 多肽，其特征在于，它包括：具有 SEQ ID NO.2 所示的氨基酸序列的多肽、或其多肽的片段、类似物或衍生物。

2、如权利要求 1 所述的多肽，其特征在于，它是具有 SEQ ID NO.2 所示的氨基酸序列的多肽或其氨基酸变异不超过 5 % 的衍生物。

3、如权利要求 2 所述的多肽，其特征在于，它是具有 SEQ ID NO.2 所示的氨基酸序列的多肽。

4、一种分离的多核苷酸，其特征在于，所述多核苷酸是选自下组：

(a) 编码具有 SEQ ID NO.2 所示氨基酸序列的多肽或其片段、类似物、衍生物的多核苷酸；

(b) 与多核苷酸(a)互补的多核苷酸；

(c) 与(a)或(b)有至少 70%相同性的多核苷酸。

5、如权利要求 4 所述的多核苷酸，其特征在于，所述多核苷酸是编码具有 SEQ ID NO.2 所示氨基酸序列的多核苷酸。

6、如权利要求 4 所述的多核苷酸，其特征在于，所述多核苷酸的序列是具有 SEQ ID NO.1 中 107-1348 位的序列或具有 SEQ ID NO.1 中 1-2925 位的序列。

7、一种含有外源多核苷酸的重组载体，其特征在于，它是由权利要求 4 所述多核苷酸与质粒、病毒或表达载体构建而成的重组载体。

8、一种含有外源多核苷酸的遗传工程化宿主细胞，其特征在于，它是选自于下列一种宿主细胞：

(a) 用权利要求 7 所述的重组载体转化或转导的宿主细胞；

(b) 用权利要求 4 所述多核苷酸转化或转导的宿主细胞。

9、一种具有 BioZFTF45 活性的多肽的制备方法，其特征在于，所述方法包括：

(a) 在适合表达 BioZFTF45 条件下，培养权利要求 8 所述的工程化宿主细胞；

(b) 从培养物中分离出具有 BioZFTF45 活性的多肽。

10、一种能与多肽结合的抗体，其特征在于，所述抗体是能与 BioZFTF45 特异性结合的抗体。

11、一类模拟或调节多肽活性或表达的化合物，其特征在于，它是模拟 BioZFTF45 的活性化合物，或促进 BioZFTF45 的活性的化合物，或拮抗 BioZFTF45 的活性的化合物，或抑制 BioZFTF45 的活性的化合物。

12、如权利要求 11 所述的化合物，其特征在于，它是 SEQ ID NO .1 所示的多核苷酸序列或其片段的反义序列。

13、一种权利要求 11 所述化合物的应用，其特征在于，所述化合物用于调节 BioZFTF45 在体内、体外活性的方法。

14、一种检测与权利要求 1 所述的 BioZFTF45 多肽相关的疾病或疾病易感性的方法，其特征在于，是通过选自下组的方法来检测相关的疾病或疾病的易感性：

(a)间接或直接检测所述多肽表达量是否异常；

(b)间接或直接检测所述多肽活性是否异常；

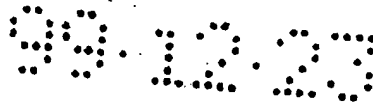
(c)直接或间接检测多核苷酸中引起所述多肽表达量或活性异常的核苷酸变异。

15、如权利要求 1 所述多肽的应用，其特征在于，它应用于筛选 BioZFTF45 的模拟物、激动剂、拮抗剂或抑制剂；或者用于肽指纹图谱鉴定。

16、如权利要求 4 所述的核酸分子的应用，其特征在于，它作为引物用于核酸扩增反应，或者作为探针用于杂交反应，或者用于制造基因芯片或微阵列。

17、一种药物组合物，其特征在于，它含有安全有效量的权利要求 1 所述的多肽和/或权利要求 4 所述的多核苷酸以及药学上可接受的载体。

18、如权利要求 1 所述的多肽的应用，其特征在于，用所述多肽制备用于用于治疗神经系统疾病、血液性恶性疾病、发育紊乱症、肿瘤、内分泌系统疾病、免疫系统疾病等疾病的药物。



## 说明书

### 新的人锌指转录因子及其编码序列

本发明属于生物技术领域和遗传工程领域，具体地说，本发明涉及了一种新的多肽——人锌指转录因子 (Novel Human Zinc Finger Transcription Factor, 简称“BioZFTF45”), 以及编码此多肽的多核苷酸序列。本发明还涉及此多核苷酸和多肽的制备方法和应用。

转录因子对基因表达的调控主要通过转录因子与特定的 DNA 序列结合、与常规转录机构的相互作用、转录因子间的相互作用来完成。根据结构基序的不同，已知的 DNA 结合蛋白可主要分为两类：含有螺旋-转角-螺旋基序的蛋白和锌指蛋白 [Kamal Chowdhury, Heidi Rohdewohld et al., Nucleic Acids Research, 1988; 16: 9995-10011]。锌指蛋白在生物体内分布非常广泛，起转录调节因子的作用，又通常称其为锌指转录因子。且已从酵母、果蝇、鼠及人等多种生物体中分离得到各种类型的锌指转录因子，形成了锌指转录因子蛋白家族。1994 年，Yamagata K. 等人从大鼠体内克隆得到了一锌指转录因子 Egr3/Pilot; 1999 年，Wani M. A. 等人从鼠中又克隆得到了一锌指转录因子 LKLF，该转录因子为 Kruppel 型锌指转录因子；同年，人们又从人体内克隆得到了一新的锌指转录因子。

锌指转录因子可按其所含锌指结构的不同来分类，如 C2H2 型、C3H 型及 C4 型等，所有的锌指结构域均是锌指转录因子与一些特殊的 DNA 结合的识别位点 [Suzuki T., Yamamoto T. et al., 1998, J Biochem, 124:389-395]。这些锌指转录因子在生物体内与基因的转录活化及抑制作用有关，调节着各种细胞的分化与增殖过程，人们在各种不同的癌组织中亦发现含有大量锌指转录因子的突变体。这些转录因子的表达异常将引发生物体内各种发育紊乱性疾病、各种肿瘤的发生、各种遗传性疾病及免疫系统疾病 [Gebelein B., Mesa K. et al., 1996, Cancer Lett, 105:225-231]。

锌指结构域在各种不同类型的锌指转录因子中均含有多个拷贝，其锌指个数不同，则功能也不同。锌指转录因子与不同长度的 DNA 的结合依赖于锌指结构的数量，多指结构可能与复合物的结合稳定性有关，而复合物是 RNA 聚合酶转录的作用位点。研究发现，许多锌指转录因子的锌指结构域相互连接区域也是高度保守的，这一区域通常含有下列序列：His-Thr-Gly-Gly-Lys-Pro-(Tyr, Phe)-X-Cys, 其中组氨酸与半胱氨酸为金属离子的结合位点，而 X 为可变氨基酸残基。这一区域对于锌指结构的形成是必需的 [Jeremy M. Berg, Annu.



Rev. Biophys. Chem, 1990, 19:405-421].

研究发现, 该蛋白家族的成员在生物体内参与调节肺、脑、脾及神经系统等多种组织的基因的表达, 其表达异常通常与生物体内一些发育紊乱性疾病、免疫系统疾病、神经系统疾病及各种肿瘤及癌症等疾病的发生相关。因此, 为治疗目的研究和开发人锌指转录因子有重要意义。

本发明的一个目的是提供分离的新的多肽——人锌指转录因子(简称为“BioZTF45”)以及其片段、类似物和衍生物。

本发明的另一个目的是提供编码该 BioZTF45 多肽的多核苷酸。

本发明的另一个目的是提供含有编码 BioZTF45 的多核苷酸的重组载体。

本发明的另一个目的是提供含有编码 BioZTF45 的多核苷酸的基因工程化宿主细胞。

本发明的另一个目的是提供生产 BioZTF45 的方法。

本发明的另一个目的是提供针对本发明的 BioZTF45 多肽的抗体。

本发明的另一个目的是提供了针对本发明 BioZTF45 多肽的模拟化合物、拮抗剂、激动剂、抑制剂。

本发明的另一个目的是提供诊断和治疗与 BioZTF45 异常相关的疾病的方法。

在本发明的第一方面, 提供新颖的分离出的人锌指转录因子 (BioZTF45), 该多肽是人源的, 它包含: 具有 SEQ ID NO: 2 氨基酸序列的多肽、或其保守性变异多肽、或其活性片段、或其活性衍生物、类似物。较佳地, 该多肽是具有 SEQ ID NO: 2 氨基酸序列的多肽或其氨基酸变异不超过 5 % 的衍生物。

在本发明的第二方面, 提供分离的编码这些多肽的多核苷酸, 该多核苷酸包含一核苷酸序列, 该核苷酸序列与选自下组的一种核苷酸序列有至少 70% 相同性: (a) 编码上述 BioZTF45 的多核苷酸; (b) 与多核苷酸(a) 互补的多核苷酸。较佳地, 该多核苷酸编码具有 SEQ ID NO: 2 所示氨基酸序列的多肽。更佳地, 该多核苷酸的序列是选自下组的一种: (a) 具有 SEQ ID NO: 1 中 107-1348 位的序列; 和(b) 具有 SEQ ID NO: 1 中 1-2925 位的序列。

在本发明的第三方面, 提供了含有上述多核苷酸的载体, 以及被该载体转化或转导的宿主细胞或者被上述多核苷酸直接转化或转导的宿主细胞。

本发明的其它方面由于本文的技术的公开, 对本领域的技术人员而言是显而易见的。



下列附图用于说明本发明的具体实施方案，而不用限定由权利要求书所界定的本发明范围。

图1是本发明的锌指转录因子(BioZFTF45)和Homo sapiens的锌指转录因子(AF049907)的氨基酸序列同源性比较图。相同氨基酸在两个序列间用单字符氨基酸表示，相似氨基酸用“+”表示。

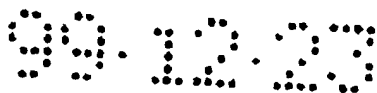
如本发明所用，“分离的”是指物质从其原始环境中分离出来(如果是天然物质，原始环境即是天然环境)。例如，活体细胞内的天然状态下的多聚核苷酸和多肽是没有分离纯化的，但同样的多聚核苷酸或多肽如从天然状态中同存在的其他物质中分开，则为分离纯化的。

如本文所用，“分离的 BioZFTF45 蛋白或多肽”是指 BioZFTF45 基本上不含天然与其相关的其它蛋白、脂类、糖类或其它物质。本领域的技术人员能用标准的蛋白质纯化技术纯化 BioZFTF45。基本上纯的多肽在非还原聚丙烯酰胺凝胶上能产生单一的主带。BioZFTF45 多肽的纯度能用氨基酸分析确定。

本发明提供了一种新的多肽——BioZFTF45多肽，其基本上是由SEQ ID NO: 2所示的氨基酸序列组成的。本发明的多肽可以是重组多肽、天然多肽、合成多肽，优选重组多肽。本发明的多肽可以是天然纯化的产物，或是化学合成的产物，或使用重组技术从原核或真核宿主(例如，细菌、酵母、高等植物、昆虫和哺乳动物细胞)中产生。根据重组生产方案所用的宿主，本发明的多肽可以是糖基化的，或可以是非糖基化的。本发明的多肽还可包括或不包括起始的甲硫氨酸残基。

本发明还包括 BioZFTF45 的片段、衍生物和类似物。如本发明所用，术语“片段”、“衍生物”和“类似物”是指基本上保持本发明天然的 BioZFTF45 相同的生物学功能或活性的多肽。本发明多肽的片段、衍生物或类似物可以是：(i)有一个或多个保守或非保守性氨基酸残基(优选保守性氨基酸残基)被取代的多肽，而这样的取代的氨基酸残基可以是也可以不是由遗传密码编码的，或(ii)在一个或多个氨基酸残基中具有取代基团的多肽，或(iii)成熟多肽与另一个化合物(比如延长多肽半衰期的化合物，例如聚乙二醇)融合所形成的多肽，或(iv)附加的氨基酸序列融合到此多肽序列而形成的多肽(如前导序列或分泌序列或用来纯化此多肽的序列或蛋白原序列)。根据本文的教导，这些片段、衍生物和类似物属于本领域熟练技术人员公知的范围之内。

本发明还提供了分离的核酸(多核苷酸)，该多核苷酸基本由编码具有 SEQ ID NO: 2 氨



基酸序列的多肽的多核苷酸组成。较佳地，本发明的多核苷酸序列具有 SEQ ID NO: 1 的核苷酸序列。

本发明的多核苷酸可以是 DNA 形式或是 RNA 形式。DNA 形式包括 cDNA、基因组 DNA 或人工合成的 DNA。DNA 可以是单链的或是双链的。DNA 可以是编码链或非编码链。编码成熟多肽的编码区序列可以与 SEQ ID NO:1 所示的编码区序列相同或者是简并的变异体。如本发明所用，“简并的变异体”在本发明中是指编码具有 SEQ ID NO:2 的蛋白质或多肽，但与 SEQ ID NO:1 所示的编码区序列有差别的核酸序列。

编码 SEQ ID NO:2 的成熟多肽的多核苷酸包括：只有成熟多肽的编码序列；成熟多肽的编码序列和各种附加编码序列；成熟多肽的编码序列(和任选的附加编码序列)以及非编码序列。

术语“编码多肽的多核苷酸”是指包括编码此多肽的多核苷酸和包括附加编码和/或非编码序列的多核苷酸。

本发明还涉及上述多核苷酸的变异体，其编码与本发明有相同的氨基酸序列的多肽或多肽的片段、类似物和衍生物。此多核苷酸的变异体可以是天然发生的等位变异体或非天然发生的变异体。这些核苷酸变异体包括取代变异体、缺失变异体和插入变异体。如本领域所知的，等位变异体是一种多核苷酸的替换形式，它可能是一个或多个核苷酸的取代、缺失或插入，但不会从实质上改变其编码的多肽的功能。

本发明还涉及与以上所描述的序列杂交的多核苷酸(两个序列之间具有至少 50%，优选具有 70%的相同性)。本发明特别涉及在严格条件下与本发明所述多核苷酸可杂交的多核苷酸。在本发明中，“严格条件”是指：(1)在较低离子强度和较高温度下的杂交和洗脱，如 0.2×SSC, 0.1%SDS, 60℃;或(2)杂交时加有变性剂，如 50%(v/v)甲酰胺, 0.1%小牛血清/0.1%Ficoll, 42℃等;或(3)仅在两条序列之间的相同性至少在 95%以上,更好是 97%以上时才发生杂交。并且，可杂交的多核苷酸编码的多肽与 SEQ ID NO: 2 所示的成熟多肽有相同的生物学功能和活性。

本发明还涉及与以上所描述的序列杂交的核酸片段。如本发明所用，“核酸片段”的长度至少含 10 个核苷酸，较好是至少 30 个核苷酸，更好是至少 50 个核苷酸，最好是至少 100 个核苷酸以上。核酸片段也可用于核酸的扩增技术(如 PCR)以确定和/或分离编码 BioZFTF45 的多核苷酸。

本发明中的多肽和多核苷酸优选以分离的形式提供，更佳地被纯化至均质。

本发明的编码 BioZFTF45 的特异的多核苷酸序列能用多种方法获得。例如，用本领域熟知的杂交技术分离多核苷酸。这些技术包括但不限于：1)用探针与基因组或 cDNA 文库杂交以检出同源的多核苷酸序列，和 2)表达文库的抗体筛选以检出具有共同结构特

征的克隆的多核苷酸片段。

本发明的 DNA 片段序列也能用下列方法获得：1) 从基因组 DNA 分离双链 DNA 序列；  
2) 化学合成 DNA 序列以获得所述多肽的双链 DNA。

上述提到的方法中，分离基因组 DNA 最不常用。DNA 序列的直接化学合成是经常选用的方法。更经常选用的方法是 cDNA 序列的分离。分离感兴趣的 cDNA 的标准方法是从高表达该基因的供体细胞分离 mRNA 并进行逆转录，形成质粒或噬菌体 cDNA 文库。提取 mRNA 的方法已有多种成熟的技术，试剂盒也可从商业途径获得(Qiagene)。而构建 cDNA 文库也是通常的方法(Sambrook, et al., Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory, New York, 1989)。还可得到商业供应的 cDNA 文库，如 Clontech 公司的不同 cDNA 文库。当结合使用聚合酶链反应技术时，即使极少的表达产物也能克隆。

可用常规方法从这些 cDNA 文库中筛选本发明的基因。这些方法包括(但不限于)：  
(1) DNA-DNA 或 DNA-RNA 杂交；(2) 标志基因功能的出现或丧失；(3) 测定 BioZFTF45 转录本的水平；(4) 通过免疫学技术或测定生物学活性，来检测基因表达的蛋白产物。上述方法可单用，也可多种方法联合应用。在第(1)种方法中，杂交所用的探针是与本发明的多核苷酸的任何一部分同源，其长度至少 10 个核苷酸，较好是至少 3 个核苷酸，更好是至少 50 个核苷酸，最好是至少 100 个核苷酸。此外，探针的长度通常在 2000 个核苷酸之内，最佳的为 1000 个核苷酸之内。此处所用的探针通常是在本发明的基因序列信息的基础上化学合成的 DNA 序列。本发明的基因本身或者片段当然可以用作探针。DNA 探针可用放射性同位素、荧光素或酶(如碱性磷酸酶)等进行标记。

在第(4)种方法中，检测 BioZFTF45 基因表达的蛋白产物可用免疫学技术如 Western 印迹法，放射免疫沉淀法，酶联免疫吸附法(ELISA)等。

应用 PCR 技术扩增 DNA/RNA 的方法(Saiki, et al. Science 1985;230:1350-1354)被优选用于获得本发明的基因。特别是很难从文库中得到全长的 cDNA 时，可优选使用 RACE 法(RACE-cDNA 末端快速扩增法)。用于 PCR 的引物可根据本文所公开的本发明的多核苷酸序列信息适当地选择，并可用常规方法合成。可用常规方法如通过凝胶电泳分离和纯化扩增的 DNA/RNA 片段。

如上所述得到的本发明的基因或其各种 DNA 片段的核苷酸序列，可用常规方法如双脱氧链终止法(Sanger et al. PNAS, 1977, 74: 5463-5467)测定。这类多核苷酸序列测定也可用商业测序试剂盒。为了获得全长的 cDNA 序列，测序需反复进行。有时需要测定多个克隆的 cDNA 序列，才能拼接成全长的 cDNA 序列。

本发明也涉及包含本发明的多核苷酸的载体，以及用本发明的载体或直接用 BioZFTF45 编码序列经基因工程产生的宿主细胞，以及经重组技术产生本发明所述多肽



的方法。

本发明中，编码 BioZFTF45 的多核苷酸序列可插入到载体中，以构成含有本发明所述多核苷酸的重组载体。术语“载体”指本领域熟知的细菌质粒、噬菌体、酵母质粒、植物细胞病毒、哺乳动物细胞病毒如腺病毒、逆转录病毒或其它载体。在本发明中适用的载体包括但不限于：在细菌中表达的基于 T7 启动子的表达载体 (Rosenberg, et al. Gene, 1987, 56:125)；在哺乳动物细胞中表达的 pMSXND 表达载体 (Lee and Nathans, J Bio Chem. 263:3521, 1988) 和在昆虫细胞中表达的来源于杆状病毒的载体。总之，只要能在宿主体内复制和稳定，任何质粒和载体都可以用于构建重组表达载体。表达载体的一个重要特征是通常含有复制起始点、启动子、标记基因和翻译调控元件。

可用本领域的技术人员熟知的方法来构建含编码 BioZFTF45 的 DNA 序列和合适的转录/翻译调控元件的表达载体。这些方法包括体外重组 DNA 技术、DNA 合成技术、体内重组技术等 (Sambrook, et al. Molecular Cloning, a Laboratory Manual, Cold Spring Harbor Laboratory. New York, 1989)。所述的 DNA 序列可有效连接到表达载体中的适当启动子上，以指导 mRNA 合成。这些启动子的代表性例子有：大肠杆菌的 lac 或 trp 启动子；λ 噬菌体的 PL 启动子；真核启动子包括 CMV 立即早期启动子、HSV 胸苷激酶启动子、早期和晚期 SV40 启动子、反转录病毒的 LTRs 和其它一些已知的可控制基因在原核细胞或真核细胞或其病毒中表达的启动子。表达载体还包括翻译起始用的核糖体结合位点和转录终止子等。在载体中插入增强子序列将会使其在高等真核细胞中的转录得到增强。增强子是 DNA 表达的顺式作用因子，通常大约有 10 到 300 个碱基对，作用于启动子以增强基因的转录。可举的例子包括在复制起始点晚期一侧的 100 到 270 个碱基对的 SV40 增强子、在复制起始点晚期一侧的多瘤增强子以及腺病毒增强子等。

此外，表达载体优选地包含一个或多个选择性标记基因，以提供用于选择转化的宿主细胞的表型性状，如真核细胞培养用的二氢叶酸还原酶、新霉素抗性以及绿色荧光蛋白 (GFP)，或用于大肠杆菌的四环素或氨苄青霉素抗性等。

本领域一般技术人员都清楚如何选择适当的载体/转录调控元件 (如启动子、增强子等) 和选择性标记基因。

本发明中，编码 BioZFTF45 的多核苷酸或含有该多核苷酸的重组载体可转化或转导入宿主细胞，以构成含有该多核苷酸或重组载体的基因工程化宿主细胞。术语“宿主细胞”指原核细胞，如细菌细胞；或是低等真核细胞，如酵母细胞；或是高等真核细胞，如哺乳动物细胞。宿主细胞的代表性例子有：大肠杆菌，链霉菌属；细菌细胞如鼠伤寒沙门氏菌；真菌细胞如酵母；植物细胞；昆虫细胞如果蝇 S2 或 Sf9；动物细胞如 CHO、COS 或 Bowes 黑素瘤细胞等。



用本发明所述的 DNA 序列或含有所述 DNA 序列的重组载体转化宿主细胞可用本领域技术人员熟知的常规技术进行。当宿主为原核生物如大肠杆菌时，能吸收 DNA 的感受态细胞可在指数生长期后收获，用  $\text{CaCl}_2$  法处理，所用的步骤在本领域众所周知。也可用  $\text{MgCl}_2$  进行。如果需要，转化也可用电穿孔的方法进行。当宿主是真核生物，可选用如下的 DNA 转染方法：磷酸钙共沉淀法，或者常规机械方法如显微注射、电穿孔、脂质体包装等。

通过常规的重组 DNA 技术 (Science, 1984; 224: 1431)，利用本发明的多核苷酸序列可用来表达或生产重组的 BioZFTF45。一般来说有以下步骤：

- (1). 用本发明的编码人 BioZFTF45 的多核苷酸(或变体)，或用含有该多核苷酸的重组表达载体转化或转导合适的宿主细胞；
- (2). 在合适的培养基中培养转化或转导的宿主细胞；
- (3). 从培养基或细胞中分离、纯化蛋白质。

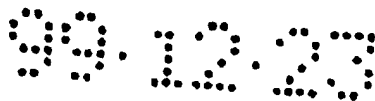
在步骤(2)中，根据所用的宿主细胞，培养中所用的培养基可选自各种常规培养基。在适于宿主细胞生长的条件下进行培养。当宿主细胞生长到适当的细胞密度后，用合适的方法(如温度转换或化学诱导)诱导选择的启动子，将细胞再培养一段时间。

在步骤(3)中，重组多肽可包被于细胞内、或在细胞膜上表达、或分泌到细胞外。如果需要，可利用其物理的、化学的和其它特性通过各种分离方法分离和纯化重组的蛋白。这些方法是本领域技术人员所熟知的。这些方法包括但不限于：常规的复性处理、蛋白沉淀剂处理(盐析方法)、离心、渗透破菌、超声波处理、超离心、分子筛层析(凝胶过滤)、吸附层析、离子交换层析、高效液相层析(HPLC)和其它各种液相层析技术及这些方法的结合。

一些含锌指结构域的锌指转录因子家族成员与以下疾病的发生有关：实体瘤如甲状腺腺瘤、子宫肌瘤；神经系统疾病如锥体外系功能异常、帕金森综合征、共济失调、神经细胞瘤、胶质细胞瘤；血液性恶性疾病如白血病、非何杰金氏淋巴瘤；发育紊乱症如 Williams 综合症，裂手裂脚症，贝魏二氏综合症；其它肿瘤如成神经细胞瘤、结肠癌、乳腺癌等。它的表达异常还将产生某些遗传性疾病，内分泌系统疾病如内分泌腺瘤，免疫系统疾病。

锌指转录因子的拮抗剂、片段、衍生物、激动剂和抑制剂可用来治疗或预防各种组织的癌变引起的疾病，这些疾病包括但(不限于)：甲状腺肿瘤、子宫肌瘤、成神经细胞瘤、结肠癌、乳腺癌、白血病、淋巴瘤、恶性组织细胞病、黑色素瘤、肉瘤、骨髓瘤、畸胎瘤等、肾上腺癌、膀胱癌、骨癌、骨髓癌、脑癌、子宫癌、胆囊癌、肝癌、肺癌、胸腺肿瘤等。

锌指转录因子的拮抗剂、片段、衍生物、激动剂和抑制剂可用来治疗或预防发育紊乱



性疾病,这些内疾病包括但(不局限于):锥体外系功能障碍、畸胎、Williams 综合症、Alagille 综合症、裂手裂脚病及贝魏二氏综合症等。

锌指转录因子的拮抗剂、片段、衍生物、激动剂和抑制剂可用来治疗或预防神经系统疾病,这些疾病包括但(不局限于):帕金森综合征、共济失调、神经细胞瘤、胶质细胞瘤等。

锌指转录因子的拮抗剂、片段、衍生物、激动剂和抑制剂可用来治疗或预防血液性恶性疾病,这些疾病包括但(不局限于):白血病、非何杰金氏淋巴瘤等。

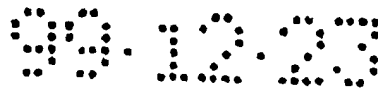
本发明也提供了筛选化合物以鉴定提高(激动剂)或阻遏(拮抗剂)BioZFTF45 的药剂的方法。激动剂提高 BioZFTF45 刺激细胞增殖等生物功能,而拮抗剂阻止和治疗与细胞过度增殖有关的紊乱如各种癌症。例如,在药物的存在下,将哺乳动物细胞与标记的 BioZFTF45 一起培养,然后测定药物提高或阻遏 BioZFTF45 作用的能力,从而鉴别出激动剂或拮抗剂。

BioZFTF45 的拮抗剂包括筛选出的抗体、化合物、受体缺失物和类似物等。BioZFTF45 的拮抗剂可以与 BioZFTF45 多肽结合并消除其功能,或是抑制该多肽的产生,或是与该多肽的活性位点结合使该多肽不能发挥生物学功能。

在筛选作为拮抗剂的化合物时,可以将 BioZFTF45 加入生物分析测定中,通过测定化合物对 BioZFTF45 和其受体之间相互作用的影响来确定化合物是否是拮抗剂。用上述筛选化合物的同样方法,可以筛选出起拮抗剂作用的受体缺失物和类似物。能与 BioZFTF45 结合的多肽分子可通过筛选由各种可能组合的氨基酸结合于固相物组成的随机多肽库而获得。筛选时,一般应对 BioZFTF45 分子进行标记。

本发明提供了用多肽,及其片段、衍生物、类似物或它们的细胞作为抗原以生产抗体的方法。这些抗体可以是多克隆抗体或单克隆抗体。本发明还提供了针对 BioZFTF45 抗原决定簇的抗体。这些抗体包括(但不限于):多克隆抗体、单克隆抗体、嵌合抗体、单链抗体、Fab 片段和 Fab 表达文库产生的片段。

多克隆抗体的生产可用 BioZFTF45 直接注射免疫动物(如家兔,小鼠,大鼠等)的方法得到。有多种佐剂可用于增强免疫反应,其中包括但不限于弗氏佐剂等。制备 BioZFTF45 的单克隆抗体的技术包括(但不限于):杂交瘤技术(Kohler and Milstein. Nature, 1975, 256:495-497),三瘤技术,人 B-细胞杂交瘤技术,EBV-杂交瘤技术等。将人恒定区和非人源的可变区结合的嵌合抗体可用已有的技术生产(Morrison et al, PNAS, 1985, 81:6851)。而已有的生产单链抗体的技术(U. S. Pat No. 496778)也可用于生



产抗 BioZFTF45 的单链抗体。

抗 BioZFTF45 的抗体可用于免疫组织化学技术中，检测活检标本中的 BioZFTF45。与 BioZFTF45 结合的单克隆抗体也可用放射性同位素标记，注入体内可跟踪其位置和分布。这种放射性标记的抗体可作为一种非创伤性诊断方法用于肿瘤细胞的定位和判断是否有转移。

抗体还可用于设计针对体内某一特殊部位的免疫毒素。如 BioZFTF45 高亲和性的单克隆抗体可与细菌或植物毒素(如白喉毒素，蓖麻蛋白，红豆碱等)共价结合。一种通常的方法是用巯基交联剂如 SPDP，攻击抗体的氨基，通过二硫键的交换，将毒素结合于抗体上，这种杂交抗体可用于杀灭 BioZFTF45 阳性的细胞。

本发明中的抗体可用于治疗或预防与 BioZFTF45 相关的疾病。给予适当剂量的抗体可以刺激或阻断 BioZFTF45 的产生或活性。

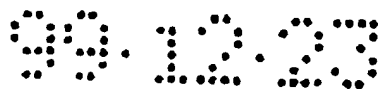
本发明还涉及定量和定位检测 BioZFTF45 水平的诊断试验方法。这些试验是本领域所熟知的，且包括 FISH 测定和放射免疫测定。试验中所检测的 BioZFTF45 水平，可以用作解释 BioZFTF45 在各种疾病中的重要性和用于诊断 BioZFTF45 起作用的疾病。

本发明的多肽还可用作肽谱分析。例如，多肽可用物理的、化学或酶进行特异性切割，并进行一维或二维或三维的凝胶电泳分析，更好的是进行质谱分析。

编码 BioZFTF45 的多核苷酸也可用于多种治疗目的。基因治疗技术可用于治疗由于 BioZFTF45 的无表达或异常/无活性表达所致的细胞增殖、发育或代谢异常。重组的基因治疗载体(如病毒载体)可设计用于表达变异的 BioZFTF45，以抑制内源性的 BioZFTF45 活性。例如，一种变异的 BioZFTF45 可以是缩短的、缺失了信号传导功能域的 BioZFTF45，虽可与下游的底物结合，但缺乏信号传导活性。因此重组的基因治疗载体可用于治疗 BioZFTF45 表达或活性异常所致的疾病。来源于病毒的表达载体如逆转录病毒、腺病毒、腺病毒相关病毒、单纯疱疹病毒、细小病毒等可用于将编码 BioZFTF45 的多核苷酸转移至细胞内。构建携带编码 BioZFTF45 的多核苷酸的重组病毒载体的方法可见于已有文献(Sambrook, et al.)。另外重组编码 BioZFTF45 的多核苷酸可包装到脂质体中然后再转移至细胞内。

多核苷酸导入组织或细胞内的方法包括：将多核苷酸直接注入到体内组织中；或在体外通过载体(如病毒、噬菌体或质粒等)先将多核苷酸导入细胞中，再将细胞移植到体内等。

抑制 BioZFTF45 mRNA 的寡核苷酸(包括反义 RNA 和 DNA)以及核酶也在本发明的范围之内。核酶是一种能特异性分解特定 RNA 的酶样 RNA 分子，其作用机制是核酶分子与互



补的靶 RNA 特异性杂交后进行核酸内切作用。反义的 RNA 和 DNA 及核酶可用已有的任何 RNA 或 DNA 合成技术获得，如固相磷酸酰胺化学合成法合成寡核苷酸的技术已广泛应用。反义 RNA 分子可通过编码该 RNA 的 DNA 序列在体外或体内转录获得。这种 DNA 序列已整合到载体的 RNA 聚合酶启动子的下游。为了增加核酸分子的稳定性，可用多种方法对其进行修饰，如增加两侧的序列长度，核糖核苷之间的连接应用磷酸硫酯键或肽键而非磷酸二酯键。

编码 BioZFTF45 的多核苷酸可用于诊断与 BioZFTF45 相关的疾病。编码 BioZFTF45 的多核苷酸可用于检测 BioZFTF45 的表达与否或在疾病状态下 BioZFTF45 的异常表达。如编码 BioZFTF45 的 DNA 序列可用于对活检标本进行杂交以判断 BioZFTF45 的表达状况。杂交技术包括 Southern 印迹法，Northern 印迹法、原位杂交等。这些技术方法都是公开的成熟技术，相关的试剂盒都可从商业途径得到。本发明的多核苷酸的一部分或全部可作为探针固定在微阵列(microarray)或 DNA 芯片(又称为“基因芯片”)上，用于分析组织中基因的差异表达分析和基因诊断。用 BioZFTF45 特异的引物进行 RNA-聚合酶链反应(RT-PCR)体外扩增也可检测 BioZFTF45 的转录产物。

检测 BioZFTF45 基因的突变也可用于诊断 BioZFTF45 相关的疾病。BioZFTF45 突变的形式包括与正常野生型 BioZFTF45 DNA 序列相比的点突变、易位、缺失、重组和其它任何异常等。可用已有的技术如 Southern 印迹法、DNA 序列分析、PCR 和原位杂交检测突变。另外，突变有可能影响蛋白的表达，因此用 Northern 印迹法、Western 印迹法可间接判断基因有无突变。

本发明的序列对染色体鉴定也是有价值的。该序列会特异性地针对某条人染色体具体位置且并可以与其杂交。目前，需要鉴定染色体上的各基因的具体位点。然而，现在只有很少的基于实际序列数据(重复多态性)的染色体标记物可用于标记染色体位置。根据本发明，为了将这些序列与疾病相关基因相关联，其重要的第一步就是将这些 DNA 序列定位于染色体上。

简而言之，根据 cDNA 制备 PCR 引物(优选 15-35bp)，可以将序列定位于染色体上。然后，将这些引物用于 PCR 筛选含各条人染色体的体细胞杂合细胞。只有那些含有相应于引物的人基因的杂合细胞会产生扩增的片段。

体细胞杂合细胞的 PCR 定位法，是将 DNA 定位到具体染色体的快捷方法。使用本发明的寡核苷酸引物，通过类似方法，可利用一组来自特定染色体的片段或大量基因组克隆而实现亚定位。可用于染色体定位的其它类似策略包括原位杂交、用标记的流式分选的染色体预筛选和杂交预选，从而构建染色体特异的 cDNA 库。

将 cDNA 克隆与中期染色体进行荧光原位杂交(FISH)，可以在一个步骤中精确地进



行染色体定位。此技术的综述，参见 Verma 等，*Human Chromosomes: a Manual of Basic Techniques*, Pergamon Press, New York (1988)。

一旦序列被定位到准确的染色体位置，此序列在染色体上的物理位置就可以与基因图数据相关联。这些数据可见于例如，*V. McKusick, Mendelian Inheritance in Man* (可通过与 Johns Hopkins University Welch Medical Library 联机获得)。然后可通过连锁分析，确定基因与业已定位到染色体区域上的疾病之间的关系。

接着，测定患病和未患病个体间的 cDNA 或基因组序列差异。如果在一些或所有的患病个体中观察到某突变，而该突变在任何正常个体中未观察到，则该突变可能是疾病的病因。比较患病和未患病个体，通常涉及首先寻找染色体中结构的变化，如从染色体水平可见的或用基于 cDNA 序列的 PCR 可检测的缺失或易位。根据目前的物理作图和基因定位技术的分辨能力，被精确定位至与疾病有关的染色体区域的 cDNA，可以是 50 至 500 个潜在致病基因间之一种 (假定 1 兆碱基作图分辨能力和每 20kb 对应于一个基因)。

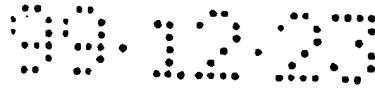
可以将本发明的多肽、多核苷酸及其模拟物、激动剂、拮抗剂和抑制剂与合适的药物载体 (药学上可接受的载体) 组合后使用。这些载体可以是水、葡萄糖、乙醇、盐类、缓冲液、甘油以及它们的组合。组合物包含安全有效量的本发明多肽或拮抗剂以及不影响药物效果的载体和赋形剂。这些组合物可以作为药物用于疾病治疗。

本发明还提供含有一种或多种容器的药盒或试剂盒，容器中装有一种或多种本发明的药用组合物成分。与这些容器一起，可以有由制造、使用或销售药品或生物制品的政府管理机构所给出的指示性提示，该提示反映出生产、使用或销售的政府管理机构许可其在人体上施用。此外，本发明的多肽可以与其它的治疗化合物结合使用。

药物组合物可以以方便的方式给药，如通过局部、静脉内、腹膜内、肌内、皮下、鼻内或皮内等的给药途径。BioZFTF45 以有效地治疗和/或预防具体的适应症的量来给药。施用于患者的 BioZFTF45 的量和剂量范围将取决于许多因素，如给药方式、待治疗者的健康条件和诊断医生的判断。

在本发明的一个实例中，提供了一种分离的多核苷酸，它编码具有 SEQ ID NO: 2 所示氨基酸序列的成熟多肽。该多核苷酸是从人胎脑组织的 cDNA 文库中发现的，多核苷酸序列全长为 2925 个碱基，其开放读框 (107-1348) 编码了 413 个氨基酸。根据氨基酸序列同源比较发现，此多肽与 *Homo sapiens* 的锌指转录因子有 53% 的同源性，由此推断本发明新的人 BioZFTF45 具有锌指转录因子基因家族相似的结构和功能。

本发明所提供的人 BioZFTF45 的 cDNA、寡聚核苷酸、多肽及抗体等，对于研究不同组织和细胞中锌指转录因子的作用、诊断锌指转录因子失调的相关性疾病、筛选抑制剂或药物治疗这些疾病有重要价值。



此外，由于本发明的人 BioZFTF45 蛋白具有源自人的天然氨基酸序列，因此，与来源于其他物种的同族蛋白相比，预计在施用于人时将具有更高的活性和/或更低的副作用(例如在人体内的免疫原性更低或没有)。

下面结合具体实施例，进一步阐述本发明。应理解，这些实施例仅用于说明本发明而不适用于限制本发明的范围。下列实施例中未注明具体条件的实验方法，通常按照常规条件如Sambrook等人，分子克隆：实验室手册(New York: Cold Spring Harbor Laboratory Press, 1989)中所述的条件，或按照制造厂商所建议的条件。

#### 实施例 1: BioZFTF45 cDNA 的克隆

用异硫氰酸胍/酚/氯仿一步法提取人胎脑总 RNA。用 Quick mRNA 分离试剂盒(Qiagen 公司产品)从总 RNA 中分离 poly(A) mRNA。2 $\mu$ g poly(A) mRNA 经逆转录形成 cDNA。用 Smart cDNA 克隆试剂盒(购自 Clontech)将 cDNA 片段定向插入到 pBSK(+)载体(Clontech 公司产品)的多克隆位点上，转化 DH5 $\alpha$ 细菌形成 cDNA 文库。用染料终止循环反应测序试剂盒(Perkin-Elmer 公司产品)和 ABI 377 自动测序仪(Perkin-Elmer 公司)测定所有克隆的 5' 和 3' 末端的序列。将测定的 cDNA 序列与已有的公共 DNA 序列数据库(Genbank)进行比较，结果发现其中一个克隆(1038b08)的 cDNA 序列为新的 DNA。通过合成一系列引物对该克隆所含的插入 cDNA 片段进行双向测定。结果表明，1038b08 克隆所含的全长 cDNA 为 2925bp(如 SEQ ID NO:1 所示)，从第 107bp 至 1348bp 有一个 1242bp 的开放阅读框架(ORF)，编码一个新的蛋白质(如 SEQ ID NO: 2 所示)。此克隆被命名为 pBS-1038b08，其编码的蛋白质命名为锌指转录因子(简称为“BioZFTF45”)。

#### 实施例2: cDNA 克隆的同源检索

将本发明的人BioZFTF45基因的核苷酸序列及其编码的蛋白序列，用Blast程序(Basic local Alignment search tool)[Altschul, SF et al. J. Mol. Biol. 1990;215:403-10]，在Genbank、Swissport等数据库进行同源检索。与本发明的人BioZFTF45基因同源性最高的基因是一种已知的Homo sapiens的锌指转录因子的基因，其编码的蛋白在Genbank的准入号为AF049907。蛋白质同源比较结果示于图1，两者高度同源，其相同性为53%；相似性为70%。这表明，本发明的新多肽具有锌指转录因子的结构和功能。

#### 实施例 3: 用 RT-PCR 方法克隆 BioZFTF45 基因



用胎脑细胞总RNA为模板，以oligo-dT为引物进行逆转录反应合成cDNA，用Qiag ne的试剂盒纯化后，用下列引物进行PCR扩增：

引物1： 5' - CATCTGGATTCATGATGGCA -3' (SEQ ID NO. 3)

引物2： 5' - CATAGGCCGAGGCGGCCGAC -3' (SEQ ID NO. 4)

引物1对应于SEQ ID NO: 1的5'端的第1bp开始的正向序列；引物2对应于SEQ ID NO: 1中的3'端反向序列。

扩增反应的条件：在50 $\mu$ l的反应体积中含有50mmol/L KCl, 10mmol/L Tris-Cl, (pH8.5), 1.5mmol/L MgCl<sub>2</sub>, 200 $\mu$ mol/L dNTP, 10pmol引物，1U的Taq DNA聚合酶(Clontech公司产品)。在PE9600型DNA热循环仪(Perkin-Elmer公司)上按下列条件反应25个周期：94 $^{\circ}$ C 30sec; 55 $^{\circ}$ C, 30sec; 72 $^{\circ}$ C 2min。在RT-PCR时同时设 $\beta$ -肌动蛋白为阳性对照和模板空白为阴性对照。扩增产物用QIAGEN公司的试剂盒纯化，用TA克隆试剂盒连接到pCR载体上(Invitrogen公司产品)。DNA序列分析结果表明，PCR产物的DNA序列与SEQ ID NO: 1所示的1-2925bp完全相同。

#### 实施例 4: Northern 印迹法分析 BioZFTF45 基因的表达：

用一步法提取总 RNA[Anal. Biochem 1987, 162, 156-159]。该法包括酸性硫氰酸胍苯酚-氯仿抽提。即用 4M 异硫氰酸胍-25mM 柠檬酸钠，0.2M 乙酸钠(pH4.0)对组织进行匀浆，加入 1 倍体积的苯酚和 1/5 体积的氯仿-异戊醇(49: 1)，混合后离心。吸出水相层，加入异丙醇(0.8 体积)并将混合物离心得到 RNA 沉淀。将得到的 RNA 沉淀用 70%乙醇洗涤，干燥并溶于水中。

用 20 $\mu$ g RNA，在含 20mM 3-(N-吗啉代)丙磺酸(pH7.0)-5mM 乙酸钠-1mM EDTA-2.2M 甲醛的 1.2%琼脂糖凝胶上进行电泳。然后转移至硝酸纤维素膜上。用 $\alpha$ -<sup>32</sup>P dATP 通过随机引物法制备 <sup>32</sup>P-标记的 DNA 探针。所用的 DNA 探针为 SEQ 1 所示的 PCR 扩增的 BioZFTF45 编码区序列(107bp 至 1348bp)。将 <sup>32</sup>P-标记的探针(约 2 $\times$ 10<sup>6</sup>cpm/ml)与转移了 RNA 的硝酸纤维素膜在溶液于 42 $^{\circ}$ C 杂交过夜，该溶液包含 50%甲酰胺-25mM KH<sub>2</sub>PO<sub>4</sub>(pH7.4)-5 $\times$ SSC-5 $\times$ Denhardt's 溶液和 200 $\mu$ g/ml 鲑精 DNA。杂交之后，将滤膜在 1 $\times$ SSC-0.1%SDS 中于 55 $^{\circ}$ C 洗 30min。然后，用 Phosphor Imager 进行分析和定量。

#### 实施例 5: 重组 BioZFTF45 的体外表达、分离和纯化

根据 SEQ ID NO: 1 的编码区序列(107-1348 位)，设计出一对特异性扩增引物，序列如下：

引物 3: 5' - CCCGGATCCATGGAGAAAGGTGGAACAT -3' (SEQ ID No 5)





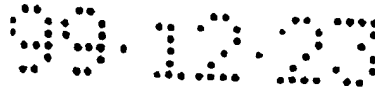
引物 4: 5' - CCCGCGGCCGCTTACCTGAGGTCAGGAGTTT -3' (SEQ ID No 6)

此两段引物的 5' 端分别含有 BamHI 和 NotI 酶切位点, 其后分别为目的基因 5' 端和 3' 端的编码序列, BamHI 和 NotI 酶切位点相应于表达载体质粒 pET-28b(+) (Novagen 公司产品, Cat.No.69865.3) 上的选择性内切酶位点。以含有全长目的基因的 pBS-1038b08 质粒为模板, 进行 PCR 反应。PCR 反应条件为: 总体积 50 $\mu$ l 中含 pBS-1038b08 质粒 10pg、引物 3 和引物 4 分别为 10pmol、Advantage polymerase Mix (Clontech 公司产品) 1 $\mu$ l。循环参数: 94  $^{\circ}$ C 20s, 60  $^{\circ}$ C 30s, 68  $^{\circ}$ C 2 min, 共 25 个循环。用 BamHI 和 NotI 分别对扩增产物和质粒 pET-28(+) 进行双酶切, 分别回收大片段, 并用 T4 连接酶连接。连接产物转化用氯化钙法大肠杆菌 DH5 $\alpha$ , 在含卡那霉素(终浓度 30 $\mu$ g/ml) 的 LB 平板培养过夜后, 用菌落 PCR 方法筛选阳性克隆, 并进行测序。挑选序列正确的阳性克隆(pET-1038b08) 用氯化钙法将重组质粒转化大肠杆菌 BL21(DE3)plySs(Novagen 公司产品)。在含卡那霉素(终浓度 30 $\mu$ g/ml) 的 LB 液体培养基中, 宿主菌 BL21(pET-1038b08) 在 37 $^{\circ}$ C 培养至对数生长期, 加入 IPTG 至终浓度 1mmol/L, 继续培养 5 小时。离心收集菌体, 经超声波破菌, 离心收集上清, 用能与 6 个组氨酸(6His-Tag) 结合的亲和层析柱 His.Bind Quick Cartridge(Novagen 公司产品) 进行层析, 得到了纯化的目的蛋白 BioZFTF45。经 SDS-PAGE 电泳, 在 45kDa 处得到一单一的条带。将该条带转移至 PVDF 膜上用 Edams 水解法进行 N-端氨基酸序列分析, 结果 N-端 15 个氨基酸与 SEQ ID NO:2 所示的 N-端 15 个氨基酸残基完全相同。

#### 实施例 6: 抗 BioZFTF45 抗体的产生

用多肽合成仪(PE 公司产品) 合成下述 BioZFTF45 特异性的多肽: Met-Glu-Lys-Gly-Gly-Asn-Ile-Gln-Leu-Glu-Ile-Pro-Asp-Phe-Ser (SEQ ID NO:7)。将该多肽分别与血蓝蛋白和牛血清白蛋白耦合成复合物, 方法参见: Avrameas, et al. *Immunochemistry*, 1969; 6:43。用 4mg 上述血蓝蛋白多肽复合物加上完全弗氏佐剂免疫家兔, 15 天后再用血蓝蛋白多肽复合物加不完全弗氏佐剂加强免疫一次。采用经 15 $\mu$ g/ml 牛血清白蛋白多肽复合物包被的滴定板做 ELISA 测定兔血清中抗体的滴度。用蛋白 A-Sepharose 从抗体阳性的家兔血清中分离总 IgG。将多肽结合于溴化氰活化的 Sepharose 4B 柱上, 用亲和层析法从总 IgG 中分离抗多肽抗体。免疫沉淀法证明纯化的抗体可特异性地与 BioZFTF45 结合。

在本发明提及的所有文献都在本申请中引用作为参考, 就如同每一篇文献被单独引用作为参考那样。此外应理解, 在阅读了本发明的上述讲授内容之后, 本领域技术人员可以对本发明作各种改动或修改, 这些等价形式同样落于本申请所附权利要求书所限定的范围。



## 序列表

### (1) 一般信息:

(i) 发明名称: 新的人锌指转录因子及其编码序列

(ii) 序列数目: 7

### (2) SEQ ID NO: 1 的信息:

#### (i) 序列特征:

(A) 长度: 2925bp

(B) 类型: 核酸

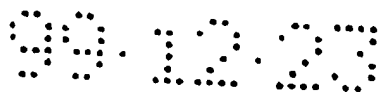
(C) 链性: 双链

(D) 拓扑结构: 线性

(ii) 分子类型: cDNA

(iii) 序列描述: SEQ ID NO: 1:

```
1 CATCTGGATTCATGATGGCATCGACAGAGGCTCTTATTGGATTAAGTTTCAAAACAAA
61 CAATAACGCTTTGGAAAAAGGGTTGAATGCACCTCAGGCAOGAOCATGGAGAAAGGTGG
121 GAACATACAATTGGAGATTCTGACTTCAGCAACTCTGTCTGAGCCATCTAAACCAGTT
181 GCGCATGCAGGGCCGTCTCTGTGATATTGTGGTCAATGTGCAAGGACAAGCTTTTGGGGC
241 TCACAAAGTGGTGCTGGCTGCCAGCTCCCCCTATTTCCGGGATCACATGTCCTTGAATGA
301 GATGAGTACAGTCTCCATTTCAAGTCAAGAACCTACTGTTTTTGAACAGCTCCTTTC
361 TTTCTGTTACGCAGGGCGGATATGCCTGCAACTGGCAGATATCATCAGCTACCTAACAGC
421 TGCCAGTTTTCTGCAATGCAGCATATTATAGACAAATGTACACAGATCCTGGAGGGCAT
481 TCATTTCAAAATTAATGTGGCTGAGGTTGAAGCAGAATTAAGTCAAACAAGGACAAAGCA
541 TCAAGAGAGACCTCCAGAGTCTCACAGGGTTACACCAATCTCAACGGCTCCCTTAGCCC
601 ACGACATAATACCCCAAAGGGAAACCGGCGAGGTCAGGTTAGTGCTGTGCTGGATATCAG
661 AGAGCTAAGTCTCCTGAGGAGTCCACCAGCCCTCAGATCATTGAACCAAGTTCTGATGT
721 AGAGAGCCGGGAGCCCATTTCTCGGATCAACCGAGCAGGACAGTGGTATGTGGAGACAGG
781 AGTGGCGGACCGTGGGGTTCGAGTGATGATGAAGTTAGAGTCTTGGAGCAGTACACAT
841 CAAAACTGAAAATCTGGAGGAGTGGCTTGGGCCTGAGAATCAGCCTTCTGGAGAAGATGG
901 GAGTAGTGCAGAGGAAGTAACAGCCATGGTGATTGATACACAGGCCATGGTTCTGTAGG
961 ACAGGAAAATTATACTTTAGGGTCTTCAGGAGCCAAGGTGGCTCGGCCAACAAAGCAGTGA
1021 AGTTGACAGATTTAGCCCCTCCGGCAGTGTGTTCCCTTGACAGAGAGACACAGAGCCAG
1081 AAGTGAGTCTCCTGGGAGAATGGATGAGCCTAAGCAACCCAGCTCCAGATGGAGTCTGG
1141 CTCTGTATCCAGGCTGGAGTGCAGTGGCGCAGCCTTGGCTCACTGCAACCTCCACCTCT
1201 CTGGTTCGAGCAATTCTCCTGTCTCAGCTTCCGAGTAGCTGGGACTACAGGTGCATGCT
1261 ACCACGCCAGCTAATTTTTGTATTTTAGTAGAGACAGGGTTTACCATATTGGTCAGG
1321 CTGGTCTCAAACCTCCTGACCTCAGGTAATCCAACCGCCTCAGCCTCCCGAAGTGCTGGGA
1381 TTACAGACATGAGCCACCGTACCCAGCCTAAACCACTGAATATTTATTGTGTGCATGGTA
1441 TAGTAATAGCCGTAAGACATTGTACAACTGTGTTTTTTAGTTTTGATAAGATGTTAGA
1501 GAAAATATTAGGAGCTTGGAGGAAAGGTAATTTGGCTCATATTATTAGGGCCCTTTTTCT
1561 CCTCATCTCTCAAATATGGTATGATAGCAATGCAGAGGAGGCTGTTGAGCAGATATAGGC
```



1621 GTGTGTCTACTAATTGACAGTATATTCTGAGAAATGCTAGACCATTTTGTTCATTGTGCAAA  
1681 TATAGAGTATATTTACACAAATTTAGATGGTATAACCTACTGCACACCTAGTTCATATGA  
1741 TATATCCCGTTCTCCTAGGCTACAAACCTGTATAGCATGTTACTGTACTGAATACTGTA  
1801 GGCAGTTACAAAGTAAATATTTGTGTATCTAAACATAAATTACTGAATATAGTAAATATT  
1861 TGTGTATCTAAACATAGAAACGATAGTGTAAAAATATTGTATTATAATCGTATGGAAGCA  
1921 CGGTTACATATGTGGTCCATCACTGACCAAAATGTTATGTGGTACATAACTGTAAATTCA  
1981 GTAATGAGTCAGACGTTACAACCTTAATGGGGGCATAAGACAGACTAACAAATACACACAA  
2041 AAAAAGTGAAGACTCTACAGTATAAACTTGAGTAATAGTGTGTGGGACTAGGATTGTT  
2101 TAAAAGGCAATGGAATAATGAAATCAGTGTCCACTTAAGGAAAGTGTGTCAAGATTTTC  
2161 TTTCTTTTTTTTTTTTTTTTTTTTTCTTTTTCGAGACAGGATCTCACTTTGTGTCAGGCG  
2221 TAGAGTGCAGTGGTATGATCATGGCTCACTGCAGCCTTGCCCTCCCTGGCTCAAGCAGTC  
2281 CCCTCACCTCAGCTCCTGAGTAGCTGGGACCACTGGCATGTGCCACCACACCCGGTTAAT  
2341 TTTTTTGTTTTTGTAGACACAGGGTTTCACTGTGTTGCCCAGGCTGATCTCGAATTCCT  
2401 GAGCTCAGTGGTGGCTCAGCCTGTAATCCCATCACTTGAAGACCAAGGTGCAGTCCTC  
2461 CCACCTGGTCTCCCAAGGGGTGAGACTACAGATGTGAGCCACCACATACAGCCTCAAGG  
2521 ATAGGGTAATATATATGCCTTAATGGATATGAAAGTTAATATGGTAAGTCTACTACCA  
2581 GTTTTTTCTCTACTACAGAGAAGTAAAAACAGGGCAAGGACTTACAGTAAGAATGAAT  
2641 ATTGATAGATTTGAAGAACTTCTGATTATGGGGGTGATTAACTGGAACACACTAATAA  
2701 AGTTACTCTTTTACTTGAAAACATTCTGTAGCTTTTTCTATGACATGAAGGTTAATATAC  
2761 TTGGCCTAATTTTTCAAGTGATTACAGTCTTTTTAACCCAAGTTTCCCACCACTCATC  
2821 AGCATTCACTGTCTCAAACAGATAGACTCTCCCTGTGTCAGAACACATTAACTATTTTC  
2881 TCTTTCTACAAAAAATACATGTGCGGCGCCTCGGCCTATG

(2) SEQ ID NO: 2 的信息:

(i) 序列特征:

(A) 长度: 413 个氨基酸

(B) 类型: 氨基酸

(D) 拓扑结构: 线性

(ii) 分子类型: 多肽

(iii) 序列描述: SEQ ID NO: 2:

1	Met	Glu	Lys	Gly	Gly	Asn	Ile	Gln	Leu	Glu	Ile	Pro	Asp	Phe	Ser
16	Asn	Ser	Val	Leu	Ser	His	Leu	Asn	Gln	Leu	Arg	Met	Gln	Gly	Arg
31	Leu	Cys	Asp	Ile	Val	Val	Asn	Val	Gln	Gly	Gln	Ala	Phe	Arg	Ala
46	His	Lys	Val	Val	Leu	Ala	Ala	Ser	Ser	Pro	Tyr	Phe	Arg	Asp	His
61	Met	Ser	Leu	Asn	Glu	Met	Ser	Thr	Val	Ser	Ile	Ser	Val	Ile	Lys
76	Asn	Pro	Thr	Val	Phe	Glu	Gln	Leu	Leu	Ser	Phe	Cys	Tyr	Ala	Gly
91	Arg	Ile	Cys	Leu	Gln	Leu	Ala	Asp	Ile	Ile	Ser	Tyr	Leu	Thr	Ala
106	Ala	Ser	Phe	Leu	Gln	Met	Gln	His	Ile	Ile	Asp	Lys	Cys	Thr	Gln
121	Ile	Leu	Glu	Gly	Ile	His	Phe	Lys	Ile	Asn	Val	Ala	Glu	Val	Glu
136	Ala	Glu	Leu	Ser	Gln	Thr	Arg	Thr	Lys	His	Gln	Glu	Arg	Pro	Pro
151	Glu	Ser	His	Arg	Val	Thr	Pro	Asn	Leu	Asn	Arg	Ser	Leu	Ser	Pro
166	Arg	His	Asn	Thr	Pro	Lys	Gly	Asn	Arg	Arg	Gly	Gln	Val	Ser	Ala

99.12.23

181	Val	Leu	Asp	Ile	Arg	Glu	Leu	Ser	Pro	Pro	Glu	Glu	Ser	Thr	Ser
196	Pro	Gln	Ile	Ile	Glu	Pro	Ser	Ser	Asp	Val	Glu	Ser	Arg	Glu	Pro
211	Ile	Leu	Arg	Ile	Asn	Arg	Ala	Gly	Gln	Trp	Tyr	Val	Glu	Thr	Gly
226	Val	Ala	Asp	Arg	Gly	Gly	Arg	Ser	Asp	Asp	Glu	Val	Arg	Val	Leu
241	Gly	Ala	Val	His	Ile	Lys	Thr	Glu	Asn	Leu	Glu	Glu	Trp	Leu	Gly
256	Pro	Glu	Asn	Gln	Pro	Ser	Gly	Glu	Asp	Gly	Ser	Ser	Ala	Glu	Glu
271	Val	Thr	Ala	Met	Val	Ile	Asp	Thr	Thr	Gly	His	Gly	Ser	Val	Gly
286	Gln	Glu	Asn	Tyr	Thr	Leu	Gly	Ser	Ser	Gly	Ala	Lys	Val	Ala	Arg
301	Pro	Thr	Ser	Ser	Glu	Val	Asp	Arg	Phe	Ser	Pro	Ser	Gly	Ser	Val
316	Val	Pro	Leu	Thr	Glu	Arg	His	Arg	Ala	Arg	Ser	Glu	Ser	Pro	Gly
331	Arg	Met	Asp	Glu	Pro	Lys	Gln	Pro	Ser	Ser	Gln	Met	Glu	Ser	Gly
346	Ser	Val	Ile	Gln	Ala	Gly	Val	Gln	Trp	Arg	Ser	Leu	Gly	Ser	Leu
361	Gln	Pro	Pro	Pro	Leu	Trp	Phe	Glu	Gln	Phe	Ser	Cys	Leu	Ser	Phe
376	Pro	Ser	Ser	Trp	Asp	Tyr	Arg	Cys	Met	Leu	Pro	Arg	Pro	Ala	Asn
391	Phe	Cys	Ile	Phe	Ser	Arg	Asp	Arg	Val	Ser	Pro	Tyr	Trp	Ser	Gly
406	Trp	Ser	Gln	Thr	Pro	Asp	Leu	Arg							

(2) SEQ ID NO: 3 的信息

(i) 序列特征

(A) 长度: 20 个碱基

(B) 类型: 核酸

(C) 链性: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: 寡核苷酸

(iii) 序列描述: SEQ ID NO : 3:

CATCTGGATTCATGATGGCA

20

(2) SEQ ID NO: 4 的信息

(i) 序列特征

(A) 长度: 20 个碱基

(B) 类型: 核酸

(C) 链性: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: 寡核苷酸

(iii) 序列描述: SEQ ID NO : 4:

CATAGGCCGAGGCGGCCGAC

20

## (2) SEQ ID NO: 5 的信息

## (i) 序列特征

(A) 长度: 29 个碱基

(B) 类型: 核酸

(C) 链性: 单链

(D) 拓扑结构: 线性

## (ii) 分子类型: 寡核苷酸

## (iii) 序列描述: SEQ ID NO : 5:

CCCGGATCCATGGAGAAAGGTGGGAACAT

29

## (2) SEQ ID NO: 6 的信息

## (i) 序列特征

(A) 长度: 31 个碱基

(B) 类型: 核酸

(C) 链性: 单链

(D) 拓扑结构: 线性

## (ii) 分子类型: 寡核苷酸

## (iii) 序列描述: SEQ ID NO : 6:

CCCGCGCCGCTTACCTGAGGTCAGGAGTTT

31

## (2) SEQ ID NO: 7 的信息:

## (i) 序列特征:

(A) 长度: 15 个氨基酸

(B) 类型: 氨基酸

(D) 拓扑结构: 线性

## (ii) 分子类型: 多肽

## (iii) 序列描述: SEQ ID NO: 7:

Met-Glu- Lys-Gly-Gly-Asn-Ile-Gln-Leu-Glu-Ile-Pro-Asp-Phe-Ser

15



99.12.23

# 说明书附图

BioZFTF45: 1 MEKGGN-IQLEIPDFSNSVLSHLNQLRMQGRLCDIVVNVQQAFAHKKVLAASSPYFRD 59  
 ME G N ++E PDFS+++L LNQ R QG+LCD+ + VQG FRAHK VLAASSPYF D  
 AF049907: 1 MEPGTNSFRVEFPDFSSTILQKLNQQRQGGQLCDVSIVVQGHIFRAHKAVLAASSPYFCD 60

BioZFTF45: 60 HMSLNEMSTVSISVIKNPTVFEQLLSFCYAGRICLQLADIISYLTAAFLQMHIIIDKCT 119  
 + L + + + NP VFE +L Y GR+ + +I+SYLTAAFLQM H++DKCT  
 AF049907: 61 QVLLKNSRRIVLPDVMNPRVFENILLSSYTGRLVMPAPEIVSYLTAAFLQMHVVDKCT 120

BioZFTF45: 120 QILEG 124  
 ++LEG  
 AF049907: 121 EVLEG 125

BioZFTF45; 本发明的人锌指转录因子

AF049907 ; Homo sapiens 的锌指转录因子

图 1

**New human zinc finger transcription factor and its code sequence**

Patent Number: CN1299825  
Publication date: 2001-06-20  
Inventor(s): MAO YUMIN (CN); XIE YI (CN)  
Applicant(s): SHANGHAI SHENGYUAN GENE DEV CO (CN)  
Requested Patent: CN1299825  
Application Number: CN19990124253 19991216  
Priority Number(s): CN19990124253 19991216  
IPC Classification: C07K14/00; C07K14/435; C07K16/00; C07K16/18; C07H21/00; C12N15/11; C12N15/63; C12N15/64  
EC Classification:  
Equivalents:

---

**Abstract**

---

The present invention discloses a novel human zinc finger transcription factor, called BioZFTF45 for short, polynucleotide for coding said polypeptide and method for producing said polypeptide by using recombination technology. Said invention also discloses the method using said polypeptide and polynucleotide to cure the diseases of nervous system disease, immunopathic malignant disease, development disturbance, tumor, endocrine system disease and immunological system disease. Said invention also discloses the antagonist for resisting said polypeptide and its therapeutic effect. Said invention also discloses the diagnostic determination method based on identification of mutation in said nucleic acid sequence and change of said polypeptide expression level, and also discloses the application of said polynucleotide coding said novel BioZFTF 45.

---

Data supplied from the esp@cenet database - I2